

CLAIMS

What is claimed is:

1. A snapshot tree structure, comprising:

a base volume storing a current user data;

a first snapshot descending from the base volume, the first snapshot being created at a first time, the first snapshot comprising:

a first table comprising a first plurality of entries corresponding to first portions of block addresses;

a second table comprising a second plurality of entries corresponding to second portions of the block addresses;

a third table comprising a third plurality of entries corresponding to third portions of the block addresses;

wherein:

one of the first plurality of entries contains a pointer to the second table;

one of the second plurality of entries contains a pointer to the third table;

one of the third plurality of entries contains a pointer to a data of the base volume at the first time before the data is modified in the base volume;

a second snapshot descending from the first snapshot, the second snapshot being created at a second time earlier than the first time.

2. The snapshot tree structure of claim 1, wherein the second snapshot comprising:

another first table comprising the first plurality of entries corresponding to the first portions of the block addresses;

another second table comprising the second plurality of entries corresponding to the second portions of the block addresses;

another third table comprising the third plurality of entries corresponding to the third portions of the block addresses;

wherein:

one of the first plurality of entries in said another first table contains a pointer to said another second table;

one of the second plurality of entries in said another second table contains a pointer to said another third table;

one of the third plurality of entries in said another third table contains a pointer to another data of the base volume at the second time before said another data is modified in the base volume.

3. The snapshot tree structure of claim 2, wherein the first and the second snapshots are read-only snapshots.
4. The snapshot tree structure of claim 2, wherein the first snapshot is a read-only snapshot and the second snapshot is a read-write snapshot.
5. A method for creating a snapshot tree structure, comprising:
 - creating a base volume;
 - at a first time, creating a first snapshot descending from the base volume, said creating a first snapshot comprising:
 - creating a first table comprising a first plurality of entries corresponding to first portions of block addresses;
 - after said creating a first snapshot, receiving a write to a data block in the base volume at a block address;
 - in response to said receiving a write, copying an original value of the data block to the first snapshot, comprising:

creating a second table comprising a second plurality of entries
corresponding to second portions of block addresses;

writing one of the first plurality of entries corresponding to a first portion
of the block address with a pointer leading to the second table;

creating a third table comprising a third plurality of entries corresponding
to third portions of block addresses;

writing one of the second plurality of entries corresponding to a second
portion of the block address with a pointer leading to the third table;

writing the original value of the data block to a new data block;

writing one of the third plurality of entries corresponding to a third
portion of the block address with a pointer leading to the new data block;

writing a new value in the data block in the base volume;

at a second time after the first time, creating a second snapshot descending from
the base volume, said creating a second snapshot comprising:

creating another first table comprising the first plurality of entries
corresponding to the first portions of block addresses;

inserting the second snapshot between the base volume and the first snapshot,
wherein the first snapshot now descends from the second snapshot.

6. The method of claim 5, further comprising, after said creating a second snapshot:

receiving another write to said data block or another data block in the base
volume, said another block having another block address;

in response to said receiving another write, copying another original value of said
data block or said another data block to the second snapshot, comprising:

creating another second table comprising the second plurality of entries
corresponding to the second portions of the block addresses;

writing, in said another first table, one of the first plurality of entries corresponding to a first address portion of said data block or said another data block with a pointer leading to said another second table;

creating another third table comprising the second plurality of entries corresponding to third portions of block addresses;

writing, in said another second table, one of the second plurality of entries corresponding to a second address portion of said data block or said another data block with a pointer leading to said another third table;

writing said another original value of said one of the data block and said another data block to another new data block;

writing, in said another third table, one of the third plurality of entries corresponding to a third address portion of said data block or said another data block with a pointer leading to said another new data block;

writing said another new value in said data block or said another data block in the base volume.

7. The method of claim 6, wherein the first and the second snapshots are read-only snapshots.